THE USE OF ENTROPY CODING IN SPEECH AND TELEVISION DIFFERENTIAL PCM SYSTEMS*

by

J. B. O'Neal, Jr.
North Carolina State University
Electrical Engineering Department

ABSTRACT

Much of the redundancy in a speech or television signal is eliminated when that signal is encoded into digital form by a differential PCM encoder. Further coding of the differential PCM output using entropy coding techniques (Huffman or Shannon-Fano coding) can result in a further increase in the signal to quantizing noise ratio of 5.6 dB without increasing the transmission rate. This conclusion is reached by comparing the performance of a DPCM system without entropy coding with one using entropy coding. The DPCM without entropy coding uses a minimum mean square error quantizer (Max quantizer) while the system with entropy coding uses an equilevel quantizer. The DPCM quantizer inputs are assumed to have a Laplacian probability density function. This is consistent with previous experiments with speech and television signals. Gaussian quantizer inputs result in only a 2.8 dB improvement for entropy coding.

*This work was supported by the Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, under contracts ~~F44620-69-C-00~~33 and F44620-70-C-0122.

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) North Carolina State University Department of Electrical Engineering Raleigh, North Carolina 27607	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

THE USE OF ENTROPY CODING IN SPEECH AND TELEVISION DIFFERENTIAL PCM SYSTEMS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

J. B. O'Neal, Jr.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
4 April 1972	11	13

8a. CONTRACT OR GRANT NO. F44620-70-C-0122 b. PROJECT NO. 9769 c. 61102F d. 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) [illegible]

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	Air Force Office of Scientific Research 1400 Wilson Boulevard (NM) Arlington, Virginia 22209

13. ABSTRACT

Much of the redundancy in a speech or television signal is eliminated when that signal is encoded into digital form by a differential PCM encoder. Further coding of the differential PCM output using entropy coding techniques (Huffman or Shannon-Fano coding) can result in a further increase in the signal to quantizing noise ratio of 5.6 dB without increasing the transmission rate. This conclusion is reached by comparing the performance of a DPCM system without entropy coding with one using entropy coding. The DPCM without entropy coding uses a minimum mean square error quantizer (Max quantizer) while the system with entropy coding uses an equilevel density function. This is consistent with previous experiments with speech and television signals. Gaussian quantizer inputs result in only a 2.8 dB improvement for entropy coding.

DD FORM 1 NOV 65 1473

Security Classification

INTRODUCTION

Differential PCM (DPCM) is an efficient way to encode highly correlated analog signals into binary form suitable for digital transmission, storage or for input to a digital computer. Based primarily on a patent by Cutler in 1952 [1], DPCM systems have been studied extensively for encoding television [2,3,4] and speech [5,6].

The quantizers used in most previous studies of DPCM have been quantizers tapered in various ways to match the amplitude density of the signal to be quantized. Such quantizers were originally studied by Panter and Dite [7] and an algorithm for their design was presented by Max [8]. Nitadori [5] and O'Neal [4] found that for speech and television, respectively, well designed DPCM systems had quantizer input signals whose probability density functions were approximately Laplacian (two-sided exponential). Paez and Glisson [9] apply Max's technique to the design of quantizers for Laplacian signals. Quantizers designed using Max's technique minimize the quantizing noise for any given number of quantizing levels and will be called Optimum Fixed N Quantizers (N = number of quantizing levels). Wood [10] and Gish and Pierce [11] have recently shown that these quantizers are not optimum whenever entropy coding is used to further encode the quantizer output but that uniform quantizers (with equally spaced levels) are very nearly optimum when used in conjunction with entropy coding. In this paper we investigate the utility of entropy coding for DPCM by comparing the performance of the two coders shown in Figure 1. The performance metric to be used is the signal to quantizing noise ratio at any given baud rate in the channel.

This study leads to the conclusion that entropy coding can increase the signal to quantizing noise ratios of DPCM for speech and television by as much as 5.6 dB when the baud rate is large. For lower baud rates the gain is somewhat less -- e.g., at a channel baud rate of six times the bandwidth a gain of 4.4 dB is possible with entropy coding. These results agree qualitatively with those of [illegible]ow [1[illegible]] who studied the statistics of a 4 bit DPCM coder for Picturephone. It is important to realize that entropy coding, which allows transmission at a rate equal to the entropy of the quantizer output, requires buffering of the bit stream and the use of variable length code words. Such a coding technique is difficult to implement and the full 5.6 dB improvement from entropy coding will be difficult to achieve with practical hardware.

The Operation of DPCM

The operation of the DPCM systems of Figure 1 can be summarized as follows:[1] The input signal $S(t)$ whose mean square value is σ^2, is sampled at twice its bandwidth W to produce a sequence of sample values $S_0, S_1, \ldots = \{S_i\}$. At the same time, the predictor makes an estimate of each sample value based on those which have preceeded it. These estimates are the sequence $\hat{S}_0, \hat{S}_1, \ldots = \{\hat{S}_i\}$. Each estimate is subtraced from the actual sample value producing a difference or error sequence $e_0, e_1, \ldots = \{e_i\}$, where $e_i = S_i - \hat{S}_i$. The prediction process in DPCM causes the members of the sequence $\{e_i\}$ to be independent. The quantizer represent each value of e_i by the nearest quantizing level. The quantizing levels are transmitted and used by the receiver to

---

[1] For a more detailed discussion of DPCM see [4].

reconstruct the original analog signal. The difference between each e_i and the quantizing level used to represent it is called quantizing error q_i. The quantizing levels actually transmitted are the sequence $\{e_i - q_i\}$ and this is the sequence used by the decoder to reconstruct the analog signal. The quantizing error shows up in the reconstructed signal as quantizing noise q(t). The decoder output is thus S(t) + q(t). When the number of quantizing levels N is large ($N \geq$ [illegible] is large enough) σ_q^2, the mean square value of q(t), is the same as the mean square value of $\{q_i\}$. The signal to quantizing noise of the DPCM system can be expressed as

$$\frac{\sigma^2}{\sigma_q^2} = \frac{\sigma^2}{\sigma_e^2}\,\frac{\sigma_e^2}{\sigma_q^2} = \frac{\sigma^2}{\sigma_e^2}\,\frac{1}{\varepsilon^2} \tag{1}$$

where $\varepsilon^2 = \frac{\sigma_q^2}{\sigma_e^2}$ is the normalized quantizing noise power and may be thought of as the quantizing noise power produced when the quantizer input $\{e_i\}$ has unit variance. The quantity ε^2 is determined by the properties of the quantizer and the probability density function of $\{e_i\}$. The quantity $\frac{\sigma^2}{\sigma_e^2}$ represents the amount by which the power of the signal can be reduced by prediction. It is a measure of the redundancy in the signal S(t) and can be determined from its statistics. This quantity $\frac{\sigma^2}{\sigma_e^2}$ will be the same for both systems in Figure 1 but ε^2 will depend on whether or not entropy coding is used. When the baud rate is large enough $\frac{\sigma^2}{\sigma_e^2}$ and ε^2 are relatively independent and can be calculated separately — this is the case we are dealing with in this paper.[2]

Relationship Between Quantizing Noise and Baud Rate

The baud rate B_a used to transmit the quantizing levels through a binary channel for the DPCM system without entropy coding -- in Figure 1(a) -- is simply $\log_2 N_a$ times the sampling rate 2W, i.e.,

$$B_a = 2W \log_2 N_a, \tag{2}$$

where N_a is the number of quantizing levels. When the input to the quantizer $\{e_i\}$ is a Laplacian signal of unit variance (i.e., one whose probability density function is $p(x) = \frac{1}{\sqrt{2}} \exp\left(-\sqrt{2}\ |x|\right)$), the mean square quantizing error approaches (see Equation (15) of Reference [4]), as N gets large,[3]

$$\varepsilon_a^2 \approx \frac{9}{2N_a^2} \,. \tag{3}$$

Using (3) in (2) gives the baud rate in terms of the quantizing noise

$$B_a \approx W \log_2 \frac{9}{2\varepsilon_a^2} \tag{4}$$

---

[2] For low baud rates where $\frac{\sigma^2}{\sigma_e^2}$ and ε^2 cannot be calculated independently, Gish [13] has developed a formula for DPCM which applies to first order Markov signals.

[3] In this paper the expression $A \approx B$ means that A is approximately equal to B and furthermore $\lim_{N\to\infty} A = B$ where N is the number of quantizing levels.

For the DPCM system with entropy coding -- in Figure 1(b) -- the baud rate in the channel is equal to the sampling rate multiplied by the entropy of the quantizer output H,

$$B_b = 2WH. \tag{5}$$

Wood [10] has shown that when N_b, the number of quantizing levels, gets large H approaches

$$H \approx H_i - \log_2 \Delta \tag{6}$$

where Δ is the step size of the uniform quantizer and H_i is the entropy of the quantizer input. The mean square quantizing error ε_b^2 is related to the step size by $\varepsilon_b^2 = \frac{1}{12}\Delta^2$ and for Laplacian input signals H_i is

$$H_i = \int_{-\infty}^{\infty} - p(x) \log_2 p(x)\, dx = \log_2 \sqrt{2}e. \tag{7}$$

Making these substitutions in (6) and (5) gives the baud rate in terms of the mean square quantizing error

$$B_b \approx W \log_2 \frac{e^2}{6\,\varepsilon_b^2} \tag{8}$$

To compare the signal to quantizing noise ratios of the two systems when their baud rates are equal we set (8) equal to (4) giving

$$\varepsilon_b^2 \approx \frac{e^2}{27}\,\varepsilon_a^2 \approx 0.272\,\varepsilon_a^2 \tag{9}$$

This is the relationship between the quantizing noise powers when the two systems are operating at the same baud rate. This equation says that when the baud rate is large a quantizing system using entropy coding can achieve a signal to quantizing noise ratio of 5.64 dB

($-10 \log \frac{e^2}{27}$) greater than a system without entropy coding. This applies to DPCM through Equation (1).

What happens at lower baud rates is shown in Figure 2. The curve for the case of entropy coding was plotted from (8). The curve for the case without entropy coding was plotted from results of Paez and Glisson [9] which show the relationship between ε_a and N_a for small N_a. Our equation (3) is inaccurate for values of N_a less than about 64. In Figure 2 the quantity $\frac{\sigma^2}{\sigma_e^2}$ in (1) was assumed to be 14 dB (i.e., $10 \log \frac{\sigma^2}{\sigma_e^2} = 14$). This means that the mean square value of the signal can be reduced by 14 dB by prediction -- this assumption is reasonably consistent with previous experiments with speech and television [4,6].

The above calculations are based on Laplacian signals because previous studies showed that for speech and television signals DPCM quantizer inputs are approximately Laplacian. These calculations are easily extended to Gaussian signals as follows. In the case of Gaussian signals (3) and (4) become, respectively, for large baud rates,

$$\varepsilon_a^2 \approx \frac{2.73}{N_a^2}, \tag{10}$$

$$B_a \approx W \log_2 \frac{\sqrt{3}\pi}{2\varepsilon_a^2} \tag{11}$$

For entropy coding (7) and (8) become, respectively

$$H_i = \log_2 \sqrt{2}\, e, \tag{12}$$

$$B_b \approx W \log_2 \frac{\pi e}{6\varepsilon_b^2} \tag{13}$$

Setting the baud rates equal, (11) = (13), gives

$$\varepsilon_b^2 \approx \frac{e}{3\sqrt{3}} \varepsilon_a^2 \approx 0.523 \varepsilon_a^2 .$$

Thus for Gaussian signals entropy coding gives a 2.81 dB ($-10 \log \frac{e}{3\sqrt{3}}$) reduction in quantizing noise over systems without it. Comparing (13) with Shannon's rate distortion function, $R = W \log_2 \frac{1}{\varepsilon^2}$, for a unit variance Gaussian signal shows that at $B_b = R$, $\varepsilon_b^2 = \frac{\pi e}{6} \varepsilon^2$. This means that for large baud rates quantizing systems with entropy coding give quantizing noise only 1.53 dB ($10 \log \frac{\pi e}{6}$) above what is possible for any encoding system.

The results of this paper apply to DPCM systems which do not change or adapt themselves to the statistics of the input signal. In adaptive systems the quantizer levels may change in accordance with certain properties of the quantizer input signal. The results of this paper may not apply to adaptive systems.

The 5.6 dB gain predicted for entropy coding assumes that the quantizer inputs are Laplacian. There is some evidence [6] that suggests that for speech, the DPCM quantizer input is more peaked than the Laplacian function and may be more accurately represented by a Gamma distribution. In this case the gain for entropy coding could be greater than the 5.6 dB calculated above.

REFERENCES

[1] C. C. Cutler, "Differential Quantization of Communication Signals," Patent No. 2,605,361, July 29, 1952 (applied for June 29, 1950).

[2] B. N. Oliver, "Efficient Coding," B.S.T.J., 31, July, 1952, pp. 724-750.

[3] C. W. Harrison, "Experiments with Linear Prediction in Television," B.S.T.J., 31, July, 1952, pp. 764-783.

[4] J. B. O'Neal, Jr., "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals," B.S.T.J., 45, pp. 689-721, May-June, 1966.

[5] K. Nitadori, "Statistical Analysis of DPCM," Electron. Commun. Japan, 48, February, 1965.

[6] R. A. McDonald, "Signal-to-Noise and Idle Channel Performance of Differential Pulse Code Modulation Systems -- Particular Applications to Voice Signals," B.S.T.J., 45, pp. 1123-1151, September, 1966.

[7] P. F. Panter and W. Dite, "Quantization Distortion in Pulse-Count Modulation with Nonuniform Spacing of Levels," Proc. IRE, 39, January, 1951, pp. 44-48.

[8] J. Max, "Quantizing for Minimum Distortion," IRE Trans. on Information Theory, vol. IT-6, pp. 7-12, March, 1960.

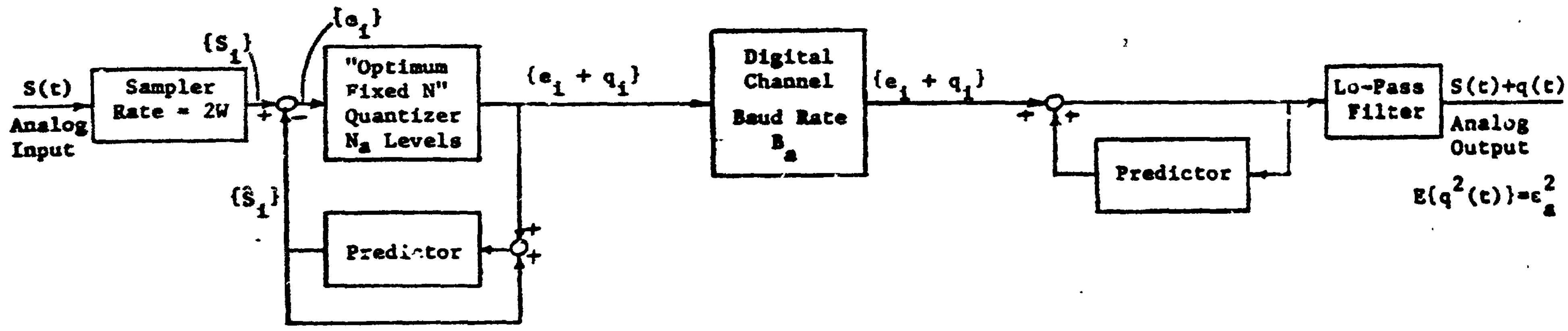
[9] M. D. Paez and T. H. Glisson, "Optimum Quantization for Laplacian Signals." to be published.

[10] R. C. Wood, "On Optimum Quantization," IEEE Trans. on Inf. Theory, Vol. IT-15, pp. 248-252, March, 1969.

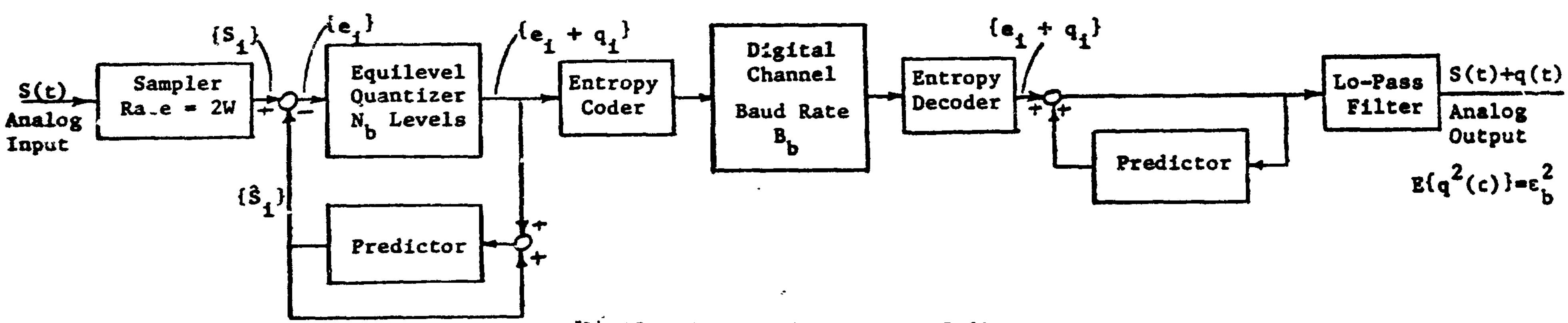
[11] H. Gish and J. N. Pierce, "Asymptotically Efficient Quantizing," IEEE Trans. on Inf. Theory, Vol. IT-14, pp. 676-683, September, 1968.

[12] M. C. Chow, "Shannon-Fano Encoding for Differentially Coded Video Telephone Signals," Proceedings of the M. J. Kelly Comm. Conf., October 5-7, 1970, paper 13-4.

[13] H. Gish, "Optimum Quantization of Random Signals," Report AD 656-042, Defense Supply Agency, Washington, D. C.



(a) DPCM System Without Entropy Coding



(b) DPCM System With Entropy Coding

Figure 1. Two differential PCM systems to be compared by setting baud rates equal, $B_a = B_b$, and computing quantizing noise powers ε_a^2 and ε_b^2.

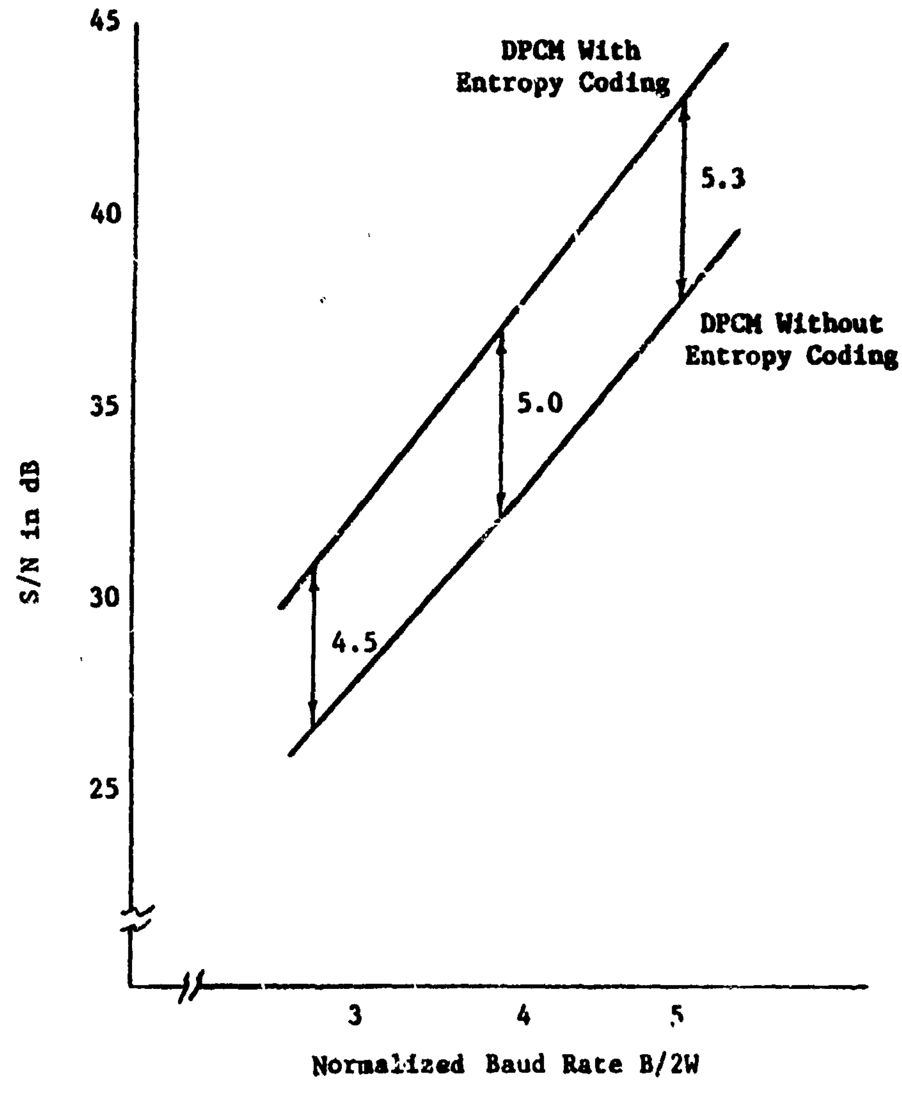


Figure 2. Signal to quantizing noise ratios S/N for the two DPCM systems in Figure 1. Both systems assume a prediction gain of 14 dB, i.e., in (1) 10 log $\sigma^2/\sigma_e^2 = 14$.